SURF

Speeded Up Robust Features

Raul Queiroz Feitosa

Objetive

This chapter presents a faster alternative to SIFT, as a scale- and rotation-invariant detector and descriptor, coined SURF.

Outline

- Overview
- Keypoint detector
- Descriptor
- Matching
- Performance Assessment
- Conclusions

Overview

Motivation

- SIFT is one of the best but slow ⇨128-D feature vectors.

SURF Characteristics

- Fast interest point **detection**
- Distinctive interest point **description**
- Speeded-up descriptor **matching**
- Invariant to common image transformations:
  - Image rotation
  - Scale changes
  - *Small* illumination change
  - *Small* change in viewpoint

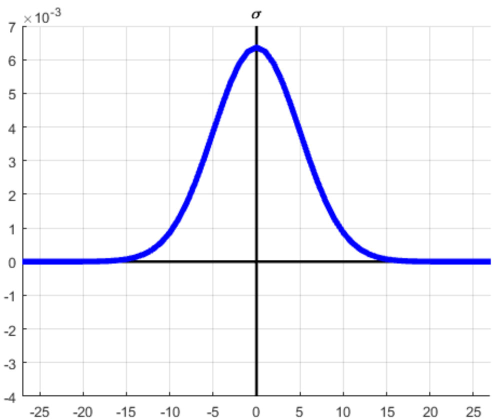
Outline

- Overview
- Keypoint detector
- Descriptor
- Matching
- Performance Assessment
- Conclusions

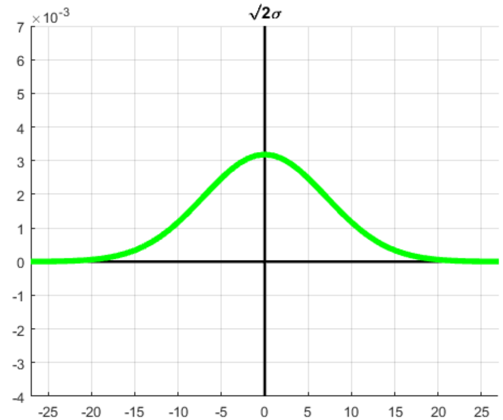
Detection

Recall that SIFT key points are given by the difference of Gaussians.

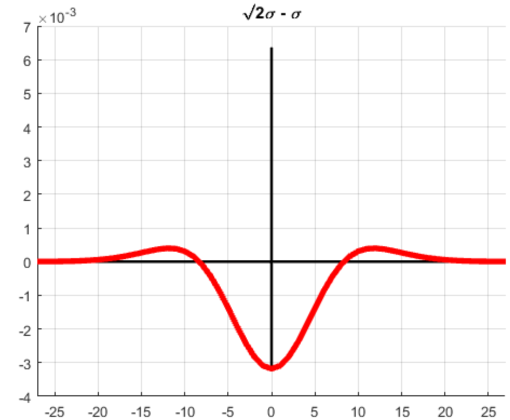
Gaussian σ



Gaussian $\sqrt{2}\sigma$

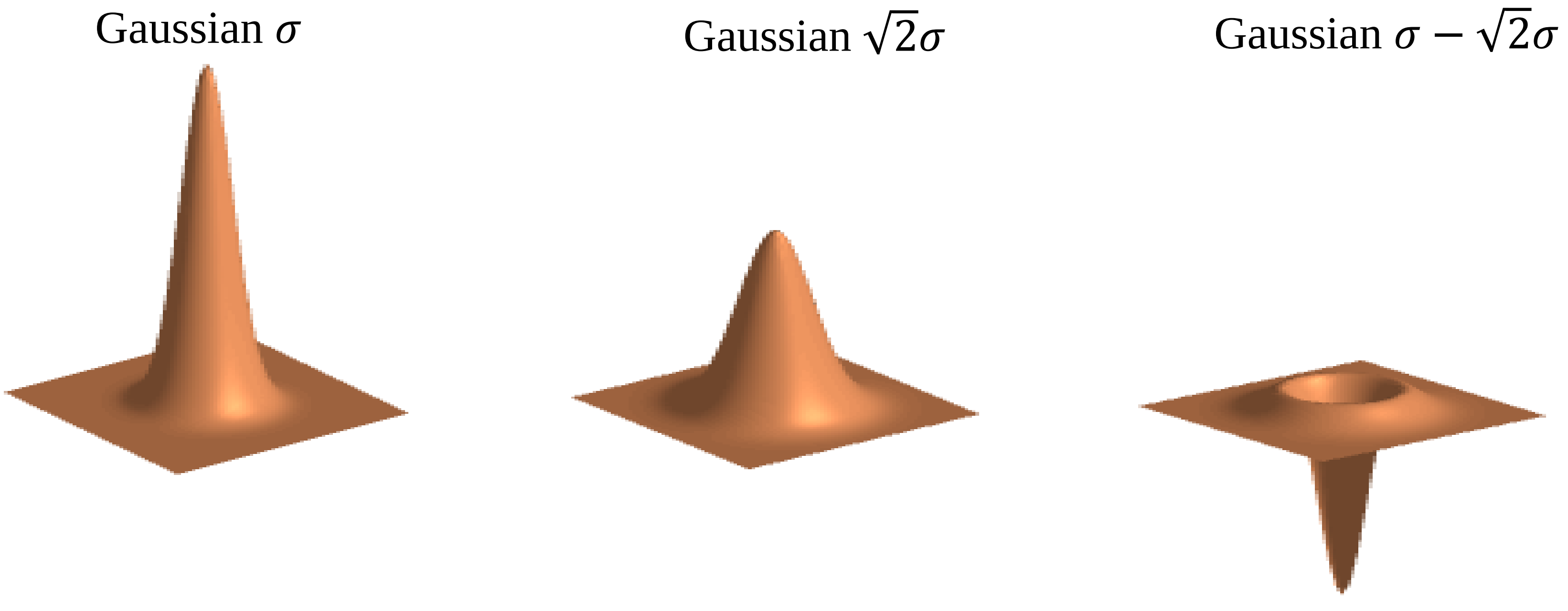


Gaussian $\sigma - \sqrt{2}\sigma$



Detection

Recall that SIFT key points are given by the difference of Gaussians.



Detection

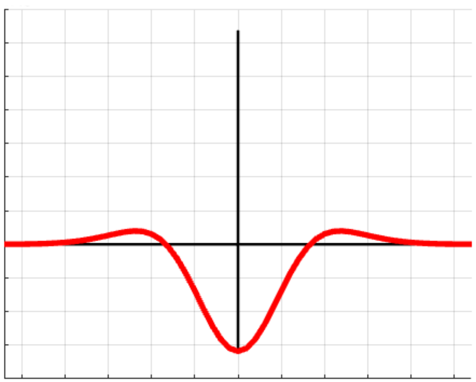
SURF relies on the determinant of the Hessian matrix

$$\mathbf{H}(s,y,\sigma) = \begin{bmatrix} L_{xx}(s,y,\sigma) & L_{xy}(s,y,\sigma) \\ L_{xy}(s,y,\sigma) & L_{yy}(s,y,\sigma) \end{bmatrix}$$

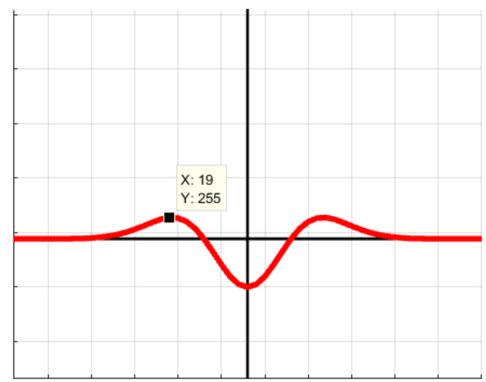
where $L_{xx}(s,y,\sigma)$ is the convolution of the Gaussian (σ) second order derivative $\frac{\partial^2}{\partial_x^2}$ with the input image in point (x,y). Similarly, $L_{yy}(s,y,\sigma)$ and $L_{xy}(s,y,\sigma)$.

Difference of Gaussians vs. Gaussian second derivative

Difference of Gaussians

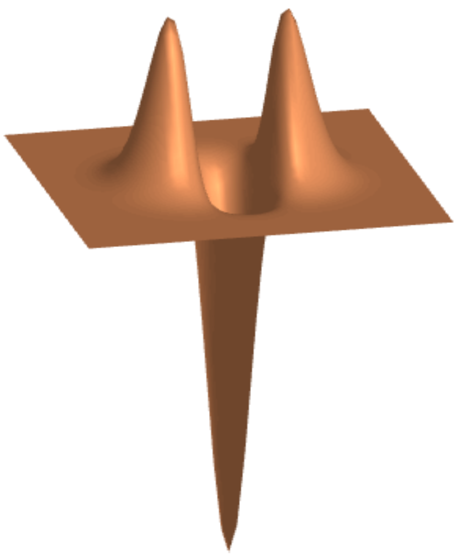


Gaussian Hessian

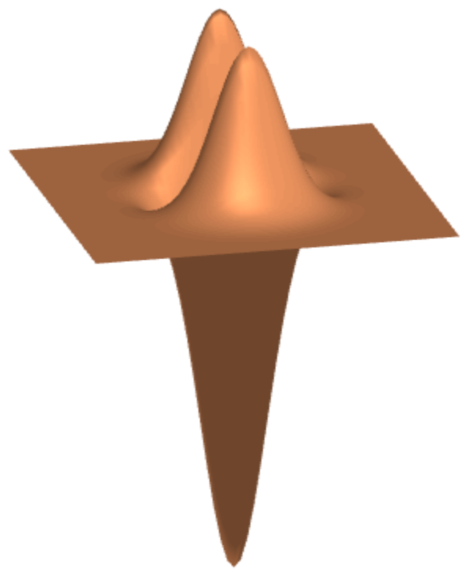


Detection

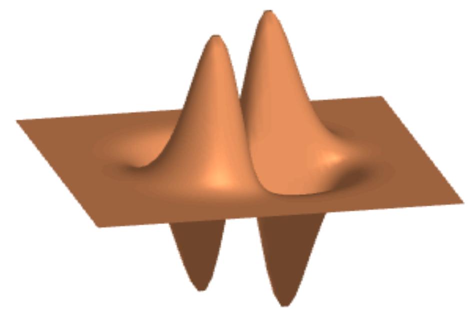
$L_{xx}(s, y, \sigma)$



$L_{yy}(s, y, \sigma)$

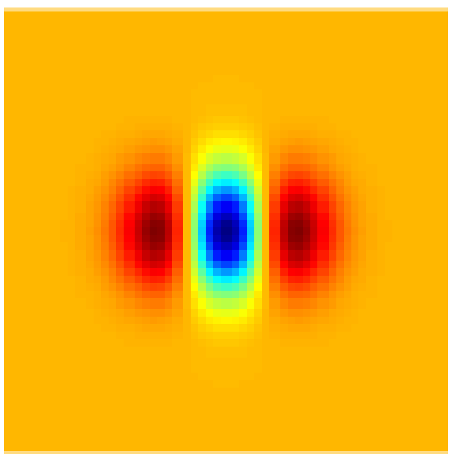


$L_{xy}(s, y, \sigma)$

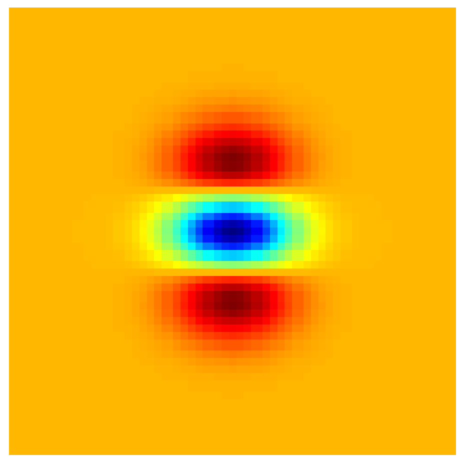


Detection

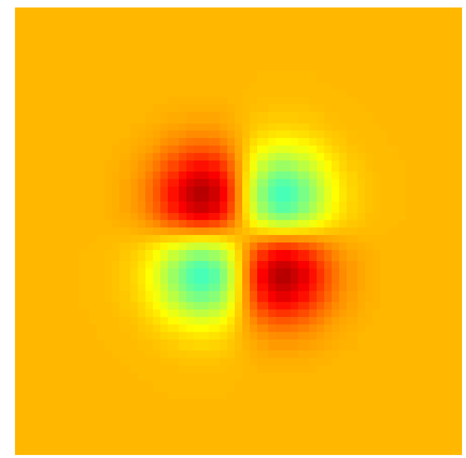
$L_{xx}(s, y, \sigma)$



$L_{yy}(s, y, \sigma)$



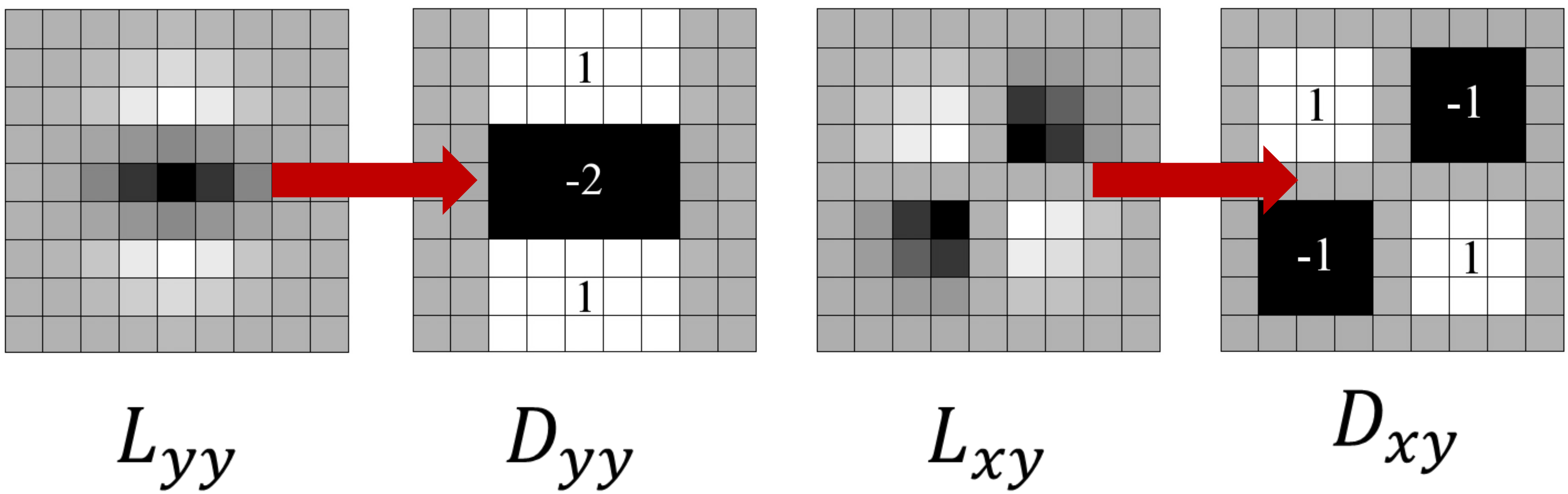
$L_{xy}(s, y, \sigma)$



Detection

Approximated second order derivatives with

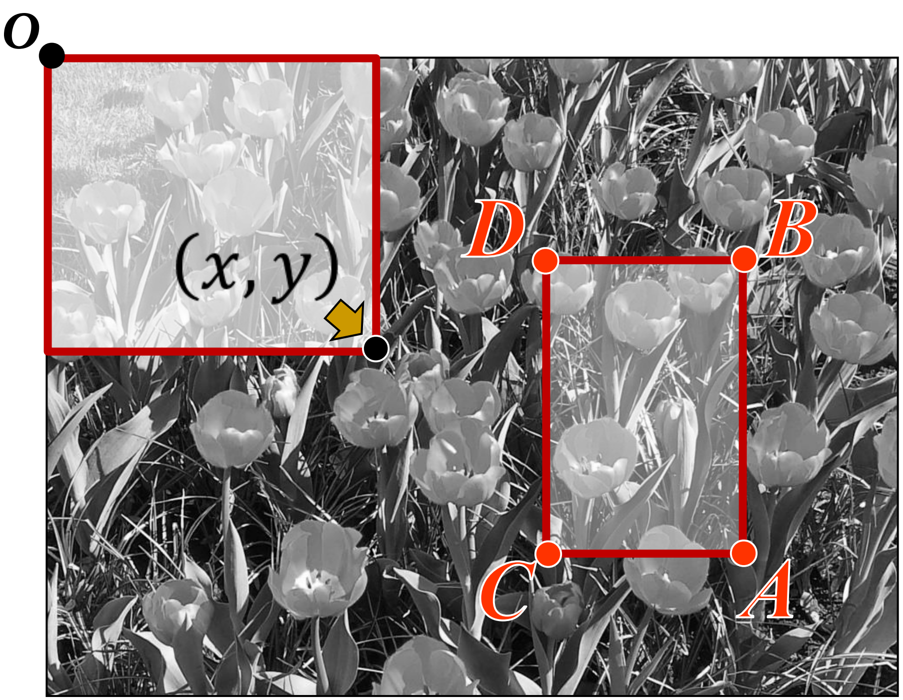
box filters



Detection

Using **integral images** for major speed up

- **Integral Image** (summed area tables) is an intermediate representation for the image and contains the **sum of gray scale pixel values of image**

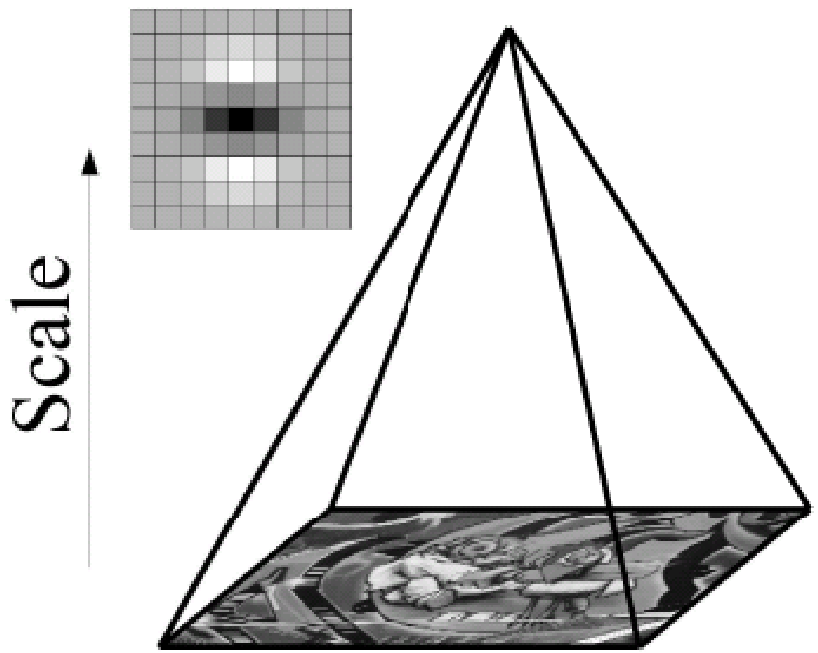


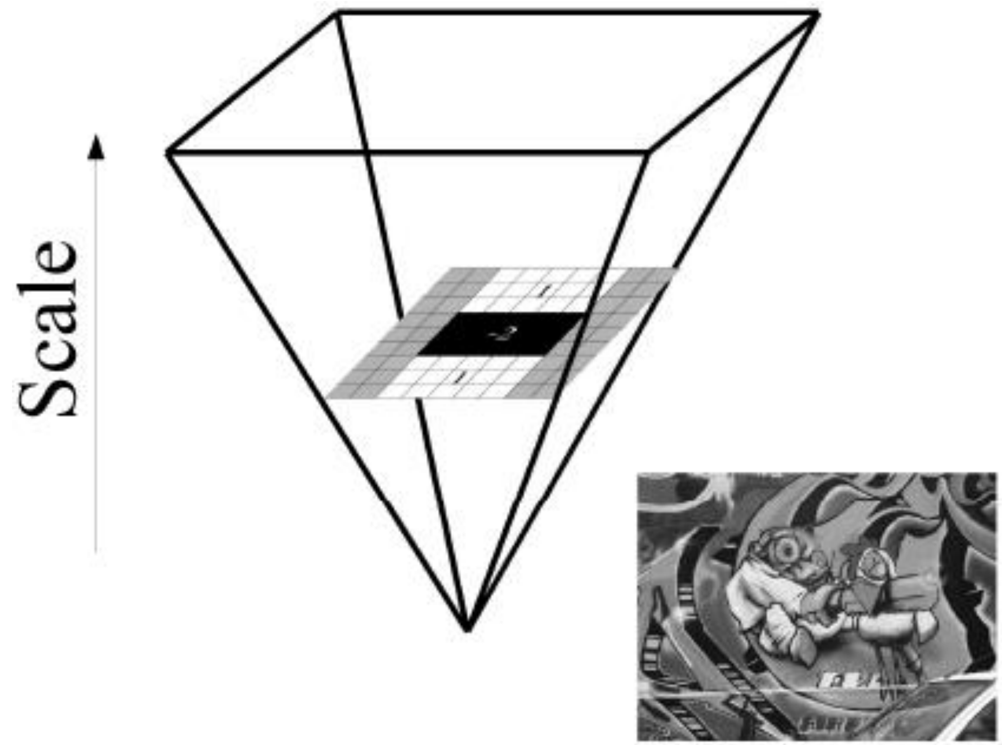
$$I_{\Sigma}(x,y) = \sum_{i=1}^{x}\sum_{j=1}^{y} \mathbf{I}(x,y)$$

Costs three additions only !!!!!!!!!

Detection

Scale analysis with **constant image size**





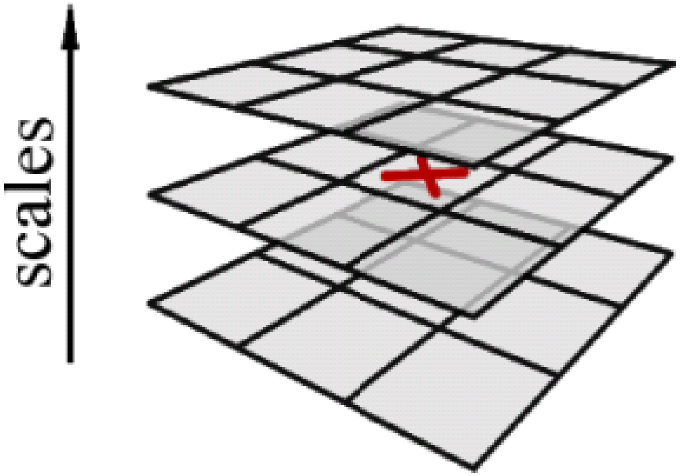
9 x 9, 15 x 15, 21 x 21, 27 x 27 → 39 x 39, 51 x 51 …

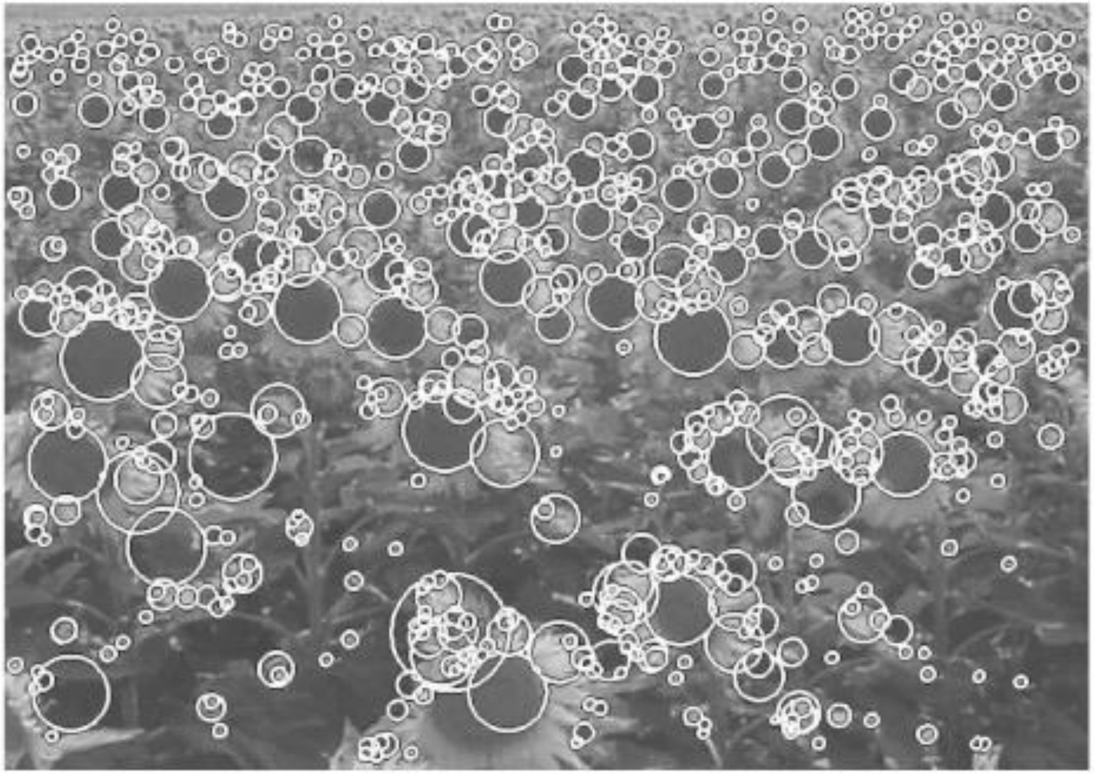
1st octave 2nd octave

Detection

Non-maximum suppression and interpolation

- Blob-like feature detector





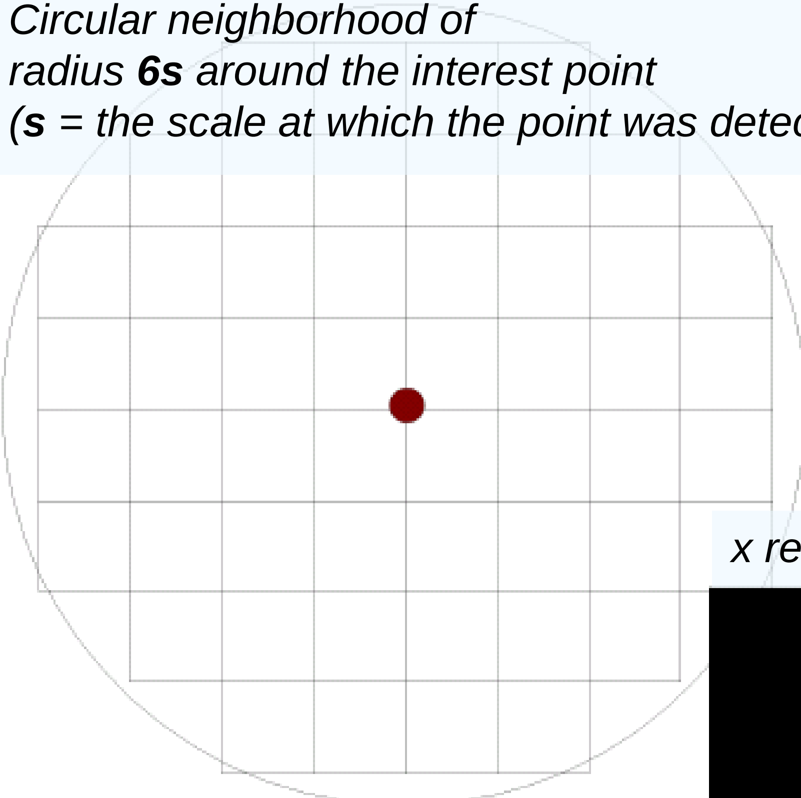
Outline

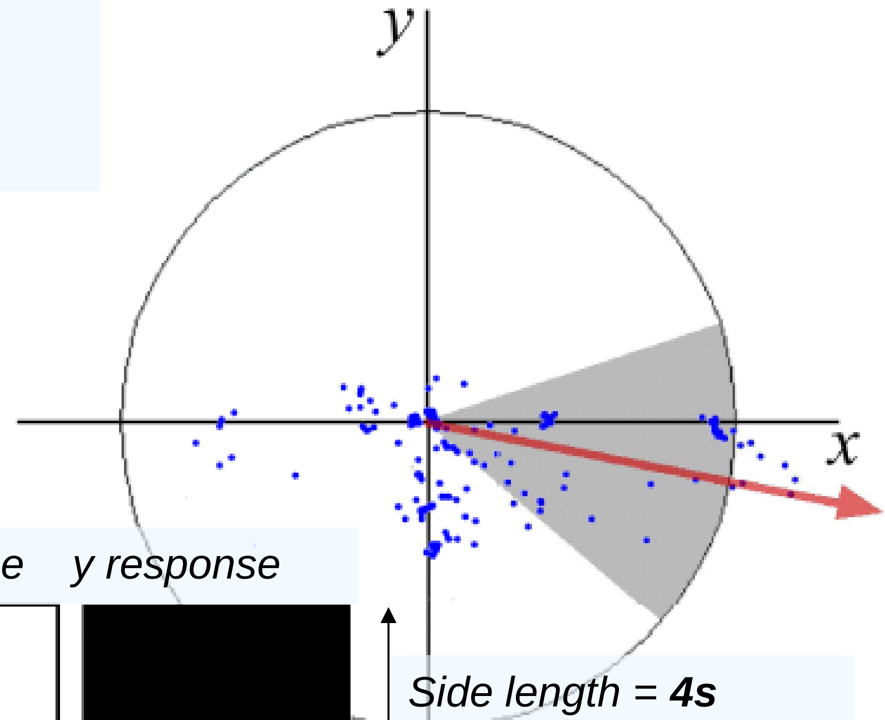
- Overview
- Keypoint detector
- Descriptor
- Matching
- Performance Assessment
- Conclusions

SURF Descriptor

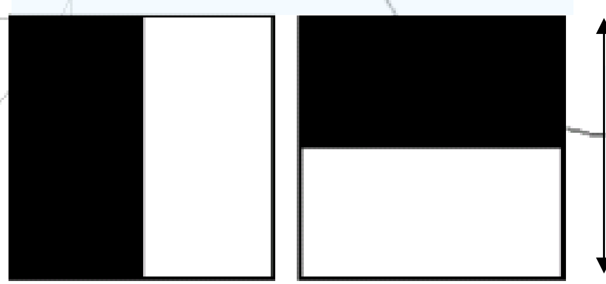
Orientation Assignment

Circular neighborhood of
*radius **6s** around the interest point*
*(**s** = the scale at which the point was detected)*





x response *y response*

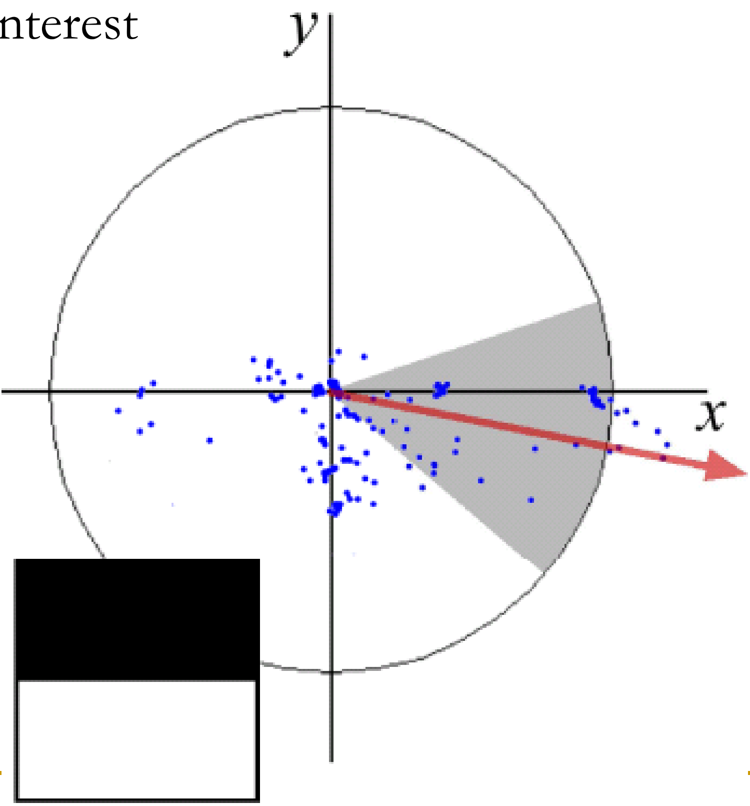


*Side length = **4s***
Costs 6 operations to compute the response

SURF Descriptor

Dominant orientation

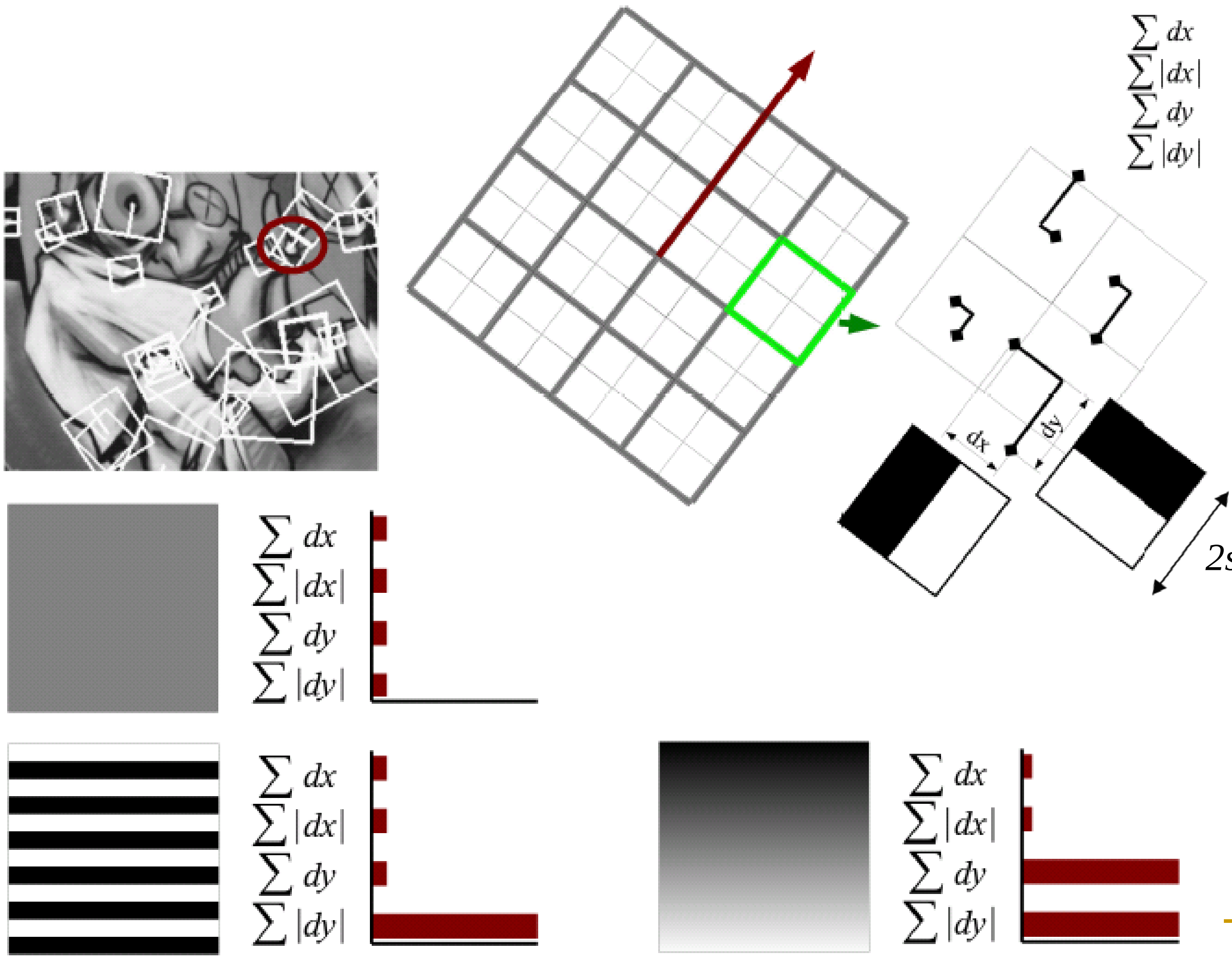
- The Haar wavelet responses are weighted with a Gaussian ($\sigma=2s$) centered at the interest point, and represented as vectors.
- Sum all responses within a sliding orientation window covering an angle of 60 degree.
- **The longest vector** is the dominant orientation
- Second longest is **ignored**



SURF Descriptor

1. Construct a square region around the interest point size = 20s, oriented according to the direction found in the previous step.
2. Split the region up into 4 x 4 square sub-regions with 5 x 5 regularly spaced sample points inside.
3. Calculate Haar wavelet response d_x and d_y (filter size = 2s) in each subregion.
4. Weight the response with a Gaussian kernel centered at the interest point.
5. Sum the response over each sub-region for d_x and d_y separately → feature vector of length 32 .
6. In order to bring in information about the polarity of the intensity changes, extract the sum of absolute value of the responses → feature vector of length 64.
7. Normalize the vector into unit length

SURF Descriptor



SURF Descriptor

SURF-128

- The sum of d_x and $|d_x|$ are computed separately for d_y<0 and d_y>0
- Similarly for the sum of d_y and $|d_y|$
- This doubles the length of a feature vector

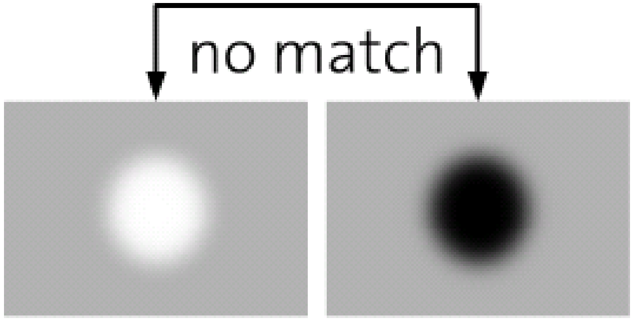
Outline

- Overview
- Keypoint detector
- Descriptor
- Matching
- Performance Assessment
- Conclusions

Matching

Fast indexing through the sign of the Laplacian for the underlying interest point

- The sign of trace of the Hessian matrix
- Trace = $L_{xx} + L_{yy}$



Either 0 or 1 (Hard thresholding, may have boundary effect …)

In the matching stage, compare features if they have the same type of contrast (sign)

Outline

- Overview
- Keypoint detector
- Descriptor
- Matching
- Performance Assessment
- Conclusions

Experimental Results

Table 1. Thresholds, number of detected points and calculation time for the detectors in our comparison. (First image of Graffiti scene, 800×640).

detector	threshold	nb of points	comp. time (msec)
Fast-Hessian	600	1418	120
Hessian-Laplace	1000	1979	650
Harris-Laplace	2500	1664	1800
DoG	default	1520	400

Table 2. Computation times for the joint detector - descriptor implementations, tested on the first image of the Graffiti sequence. The thresholds are adapted in order to detect the same number of interest points for all methods. These relative speeds are also representative for other images.

	U-SURF	SURF	SURF-128	SIFT
time (ms):	255	354	391	1036

Outline

- Overview
- Keypoint detector
- Descriptor
- Matching
- Performance Assessment
- Conclusions

Conclusion

1. **SURF describes image faster than SIFT by 3-5 times.**

2. **SURF is not as well as SIFT on invariance to**
   - illumination change and
   - viewpoint change.

SURF Reference

Paper

BAY, H.; ESS, A.; TUYTELAARS, T.; VAN GOOL, L. Speeded-up robust features (SURF). Journal of Computer Vision and Image Understanding, v.110, n.3, p. 346-359, 2008.

Code available for download

http://www.mathworks.com/matlabcentral/fileexchange/28300

Next Topic

Cameras